Decision matrix for prioritising services: first steps towards full-scale impact analysis of a public agency

Petri Mononen*, Pekka Leviäkangas and Harri Haapasalo

Industrial Engineering and Management,
University of Oulu,
P.O. Box 4610, 90014, Finland
Email: petri.mononen@oulu.fi.
Email: pekka.leviakangas@oulu.fi
Email: harri.haapasalo@oulu.fi
*Corresponding author

Abstract: Evaluating quantitative socio-economic cost-benefit level impacts of public administrations/agencies is urgently called for but still a thinly charted out territory in science. This paper presents the first steps in a process of prioritising services of a public organisation via a case study of a Finnish agency. The prioritisation is made in order to enable later analysis of the effectiveness of the agency and the evaluation of the impact of its services. The main contribution of this article is in laying out the first steps of a novel and normative process for identification of the services that are most relevant in this respect. The process entails four iterative steps: 1) service catalogue mapping; 2) design of evaluation criteria; 3) compilation of decision matrix; 4) multiple criteria decision analysis (MCDA) – and recommendations thereof. Steps 1–3 are discussed in this article in detail, the end result being a decision matrix for MCDA.

Keywords: public agency; administration; evaluation criteria; decision matrix; performance management; policy objective; impact analysis.

Reference to this paper should be made as follows: Mononen, P., Leviäkangas, P. and Haapasalo, H. (2017) 'Decision matrix for prioritising services: first steps towards full-scale impact analysis of a public agency', *Int. J. Public Sector Performance Management*, Vol. 3, No. 1, pp.1–17.

Biographical notes: Petri Mononen works as a Research Services Manager at University of Oulu, Faculty of Information Technology and Electrical Engineering. He is also a doctoral student at University of Oulu Technical Faculty, Industrial Engineering and Management. Previously he worked as Senior Scientist at Technical Research Centre of Finland VTT. Prior to joining University of Oulu and VTT, he worked as Team Leader, member of company management group and Project Manager in Liidea Ltd. Before that he has been employed by Mott MacDonald Group, UK, Jaakko Pöyry Group and City of Oulu. He is a member of the Project Management Association Finland and holds IPMA level C certificate. His main expertise areas include cost-benefit and impact evaluation, research funding grant writing, large scale international project lead/coordination, ITS in general, nomadic devices in transport, evaluation frameworks for information services and transportation system analysis.

Pekka Leviäkangas is a Research Professor and has worked in management and expert positions as a civil servant, business manager, expert and researcher. He is a Principal Scientist in Technical Research Centre of Finland and Research Professor at the University of Oulu in Industrial Engineering and Management. He has acted as Vice-President of Jaakko Pöyry Group subsidiary, Corporate Analyst of Finnish Railways (VR-Group Ltd), R&D Manager of Finnish National Road Administration's South-Eastern region, and private consultant. He is an Adjunct Professor of Tampere University of Technology in logistics and business information. His own research covers financial and socio-economic analysis, innovation management, value analysis, service sciences, project finance, investment, restructuring issues and new technology deployment, particularly in transport sector, but also in other fields such as bio-economy, climate change, meteorology, education, construction and infrastructure management.

Harri Haapasalo is a Professor of Product Management and Head of Unit at Industrial Engineering and Management, University of Oulu in Finland. He has two main areas of research; one in product management, product development and second one in area of complex systems management, lean construction and business models. He has been very active in obtaining research projects, and active in journal publications. His list of publications contains more than 200 international items. He has supervised more than 20 doctoral theses and been external examiner for more than 20 doctoral dissertations.

1 Introduction

Several economies are facing challenges in providing public services. Resources for public service production do not seem to grow in same slope. Challenges are different in different economies and so are the solution avenues. Finnish Transport Safety Agency (later referred also to as 'trafi' or 'the agency') is a good example of these phenomena, majority of the services are seen important and necessary, but there is no overall understanding of which services are absolutely necessary and which are more 'nice-to-have' services. Scarcity of public resources calls for more effective managerial grasp also in public service management. Overall the means for solving this need attention in research.

There is a long history in research in evaluating the performance of businesses, projects (Proost et al., 2014) and programs (Fuller et al., 2013). However, evaluating quantitative impacts of entire public bodies (e.g., administrations, agencies, ministries) on socio-economic benefit to cost (B/C) level is neither well-covered nor thoroughly charted out territory. In reviewing availability of relevant literature on the subject, promising material on headline level exists (e.g., European Environment Agency/COWI, 2013). However, on closer examination it turns out that regardless of thorough analysis from several viewpoints, the evaluation has not taken place in quantitative (monetary or otherwise) terms. Instead, the evaluation has been either focusing on qualitative or descriptive analysis, be it may that some quantitative attributes have been addressed. These analyses and indicators often deal with matters such as:

- public service processing, throughput and output efficiency (e.g., person years, number of permits handled, number of reports, etc.)
- evaluation of government's, other stakeholders' or clients' (direct service recipients' or the general public's) image or perceptions
- opinions on the acceptability of administration's brand value or other types of performance or 'goodness' indicators for the administration in question.

As such, these indicators often may manifest themselves as numbers on various scales, but that still does not deem them quantitative in their core nature. In fact, these kinds of indicators and metrics do not easily – if at all – directly relate to nor translate to quantitative socio-economic impacts.

When starting to approach the impact evaluation of a public body, some limitations will inherently apply. First, the direct and indirect impacts are complex and a building complete picture of the impacts and mechanisms leading to them is doable only to the extent of available resources. Secondly, even with extensive resources, the complexities and multidimensional dependencies between services or actions and direct and indirect impacts are always debatable since the causalities are far from being linear and furthermore, bound to be utmost contextual.

The end goal of an evaluation is to help increase the effectiveness of the agency as a whole. This could be in the best case measured in cost-benefit terms. Another supporting goal is to pinpoint whether some activities actually operate with a net loss benefit to cost ratio (i.e., $B/C < 1$) and hence should be either discarded or modified to enhance them. The background for the need to perform such evaluations stems partly from the recent global and national economic turmoil and downturn and the resulting increase in pressures to further decrease public expenditure and partly from the general trends of increasing the competitiveness of the nation by making its public sector as value-adding as possible. From the point of view of ever present aspiration of improved national economics and competitiveness it is nevertheless beneficial to be aiming at reducing and getting rid of the expenditure with negative or even neutral productivity. At the same time the expenditure that yields overall positive productivity and impact needs to be maintained and supported. Thus the aspiration of this research and alike will be provision of decision support for the most efficient possible use of scarce and ever scarcer resources.

2 The case subject in short

Since 2010 Finnish Transport Safety Agency has had a responsibility of maintaining and developing the safety of the transport system, promoting environmentally friendly transport solutions and is responsible for transport system regulatory duties in general. The responsibilities stem from the associated national laws and statutes (Finnish Ministry of Justice, 2009) The practical tasks performed by Trafi include issuing permits, regulations, approvals and decisions and preparing legal rules regarding the transport sector; arranging examinations, handling transport sector taxation and registration, and providing reliable information services; overseeing the transport market as well as compliance with rules and regulations governing the transport system; participating in international cooperation; ensuring the functionality of the transport system in normal

conditions, and also its resiliency in regard abnormal or emergency conditions when normal operation is disrupted; creating opportunities for innovative development of intelligent transport; and informing the public of transport-related choices. Trafi responsibilities cover the full range of transport modes, i.e., the whole transportation system. Non-motorised traffic is also included, but to a lesser extent – there are other administrations at place in Finland that have more substantial responsibilities in regard light traffic and vulnerable transport users.

Trafi personnel totalled to 494 person years in 2011 (Trafi, 2012). Trafi is organised into four branches: regulation and development, licences and approvals, oversight/enforcement and data resources; and into three processes (functions): strategy, communications and administration. Trafi additionally has mode specific directors for four transport modes, namely aviation, maritime, railways and road transport. Geographically Trafi operates in three Finnish cities, the main office is in Helsinki and additionally there are two branch offices, in Lappeenranta and Rovaniemi. Trafi also has several mini-offices located in port cities.

The total cost for Trafi operations is 142.6 million Euros annually (Trafi, 2012). This consists of costs both in the chargeable operations and un-chargeable operations. The annual budget incomes 142 million Euros main elements include governmental budgetary input (35 million Euros), incomes from chargeable activity (94 million Euros) and transfers from previous year (13 million Euros). Trafi annual turnover (annual monetary flow) also includes 765 million Euros in fiscal incomes mainly coming from governmental vehicle related taxes.

Currently the performance and the level of the agency achieving its goals are managed/controlled via a wage settlement between Finnish Ministry of Transport and Trafi (MTCF, 2012). There currently are increasing national pressures to enrich both the goal-setting and the control of those goals towards more quantitative direction, and towards actual socio-economic impacts in particular (Finnish Ministry of Finance, 2012; Finnish National Audit Office VTV, 2013). The current wage settlement (and alike, concerning other agencies across the government) do not so far respond to these pressures well. Instead, it focuses on indicators such as discussed above in previous chapter. The absence of quantitative metrics is not due to not wanting to have such metrics within the settlement but due to the fact that these metrics do not exist as of yet. Another origin of performance objectives are, though by and large through the abovementioned wage settlement conveyed through European Union transport policy objectives (European Commission, 2010, 2011). Some objectives are also channelled through Union's environmental policies.

Trafi has three main elements comprising its objectives:

1 improving traffic safety

2 mitigation of environmental impacts (mainly related to climate change and carbon emissions)

3 increasing the reliability of the transport system.

Safety objectives have a decades' long history whereas the environmental issues and most recently the underlining of the reliability of the transport system have constantly been increasing their relative importance.

3 Scope and aims

This paper presents one viable approach and process to solve the main part of the above described dilemma namely the selection of what to evaluate and what to leave out – i.e., what is reasonable and possible to be evaluated and what not. In other words, this is a presentation of a viable method to select an appropriate sample for impact evaluation research on a public agency. The paper is not arguing that the presented method is the only viable process of performing the abovementioned task of prioritising. The prioritising phase covers only the first early steps of the entire impact evaluation research entity. This article does not go into the actual impact evaluation process of the prioritised subset of services.

Public agencies and their services are under greater scrutiny than ever before. First reason for this is the trend of *new public management* (NPM) that has taken rapidly advancing steps across the globe, in countries such as Canada, New Zealand, Australia, USA, but also in Europe (Haque, 2004; Gruening, 2001). The ideological shift towards more market-oriented economy also with regard to formerly public services is often dated to Margaret Thatcher's governments in the UK. The ideas of new models of public governance brought also along the metrics of private sector management, meaning that the outputs and effectiveness of public services needed to be assessed properly in order to facilitate efficient management control. These directions were adopted by the Finnish governments since the late 1980's, when the first waves of privatisation and commercialisation of former administrative and public functions took place. Since then Finland's governments have persistently followed this path, and restructuring of public services have been carried out rather systematically (Finnish Ministry of Justice, 2009; Finnish Ministry of Finance, 2012; Finnish National Audit Office VTV, 2013). The reforms or re-structuring efforts have brought with them many performance management tools and principles, the applicability and effectiveness of which is still under debate (see e.g., Maugeri and Metzger, 2013).

The transport sector has been in the forefront of the renewal processes of different administration sectors. The Finnish Transport Agency is a fairly new entity in the transport sector governance architecture. The historical evolution of Trafi to its present form has been described in Mononen et al. (2014).

While the restructuring of administrative architecture has been systematic, it has also brought along several challenges, such as:

- difficulty to assess efficiency and effectiveness of public services as the structures – including the cost structures – have been changed and there is no historical track record upon which to build management control of costs and effectiveness
- drawing clear-cut lines between fully public, semi-public and private (or to-be-private) services in order to maximise the value for money of services provided to citizens and other customers, such as private sector companies and other sectors of public administration
- preventing *de facto* non-value adding services to exist, i.e., services that have been a part of the former public administration service architecture but for which the demand is low and which could be supplied via alternative channels or suppliers, or even cancelled entirely

- setting effective and fair performance targets for the operational management of the agency.

Hence the evaluation of any public agency's services would need to serve the purpose of providing administrations' management the information on how well the tasks and missions of the agency are fulfilled in terms of increasing socio-economic well-being and adding to the value of different functions of the society. Without this the management is not in fact doing its job in making sure that agencies and other public bodies are functioning properly, because the management lacks relevant information on the levels of performance. Partly this management control is associated with the need of continuous change, as pointed out by the OECD (OECD, 2010).

The prioritising of Trafi's services builds on the need to exercise responsible management control. The services that are after the prioritisation analysis regarded as most relevant, are later being exposed to full benefit-cost and cost-effectiveness analysis. Prior the research at hand, some traffic safety administration evaluation has been carried out by the National Audit Office (5) and Ministry of Transport and Communications Finland (MTCF, 2013). These evaluation reports dealt mainly with roles and responsibilities of country's different traffic safety actors and with Trafi's internal efficiency. While the reports excluded in-depth evaluation of Trafi's services they recognised the need to assess the benefits and costs of different functions of Trafi.

4 Research process for service prioritisation

In an ongoing research project the research subject for the impact evaluation is Trafi. Selected elements of this particular research are used in this article as a case example.

The choice of analysis method had to be considered at all times during the design of the research data format, structure and collection method design; and vice-versa, the capabilities and restrictions of all of the available analysis methods needed to be kept in mind at all times in the data structure design and collection phases. This iterative procedure and open-endedness has to be in place to avoid further mismatch between the collected data and the analysis method of choice. A flowchart illustrating the iterative nature of the prioritisation process and its research design phasing is presented in Figure 1.

4.1 Service catalogue mapping

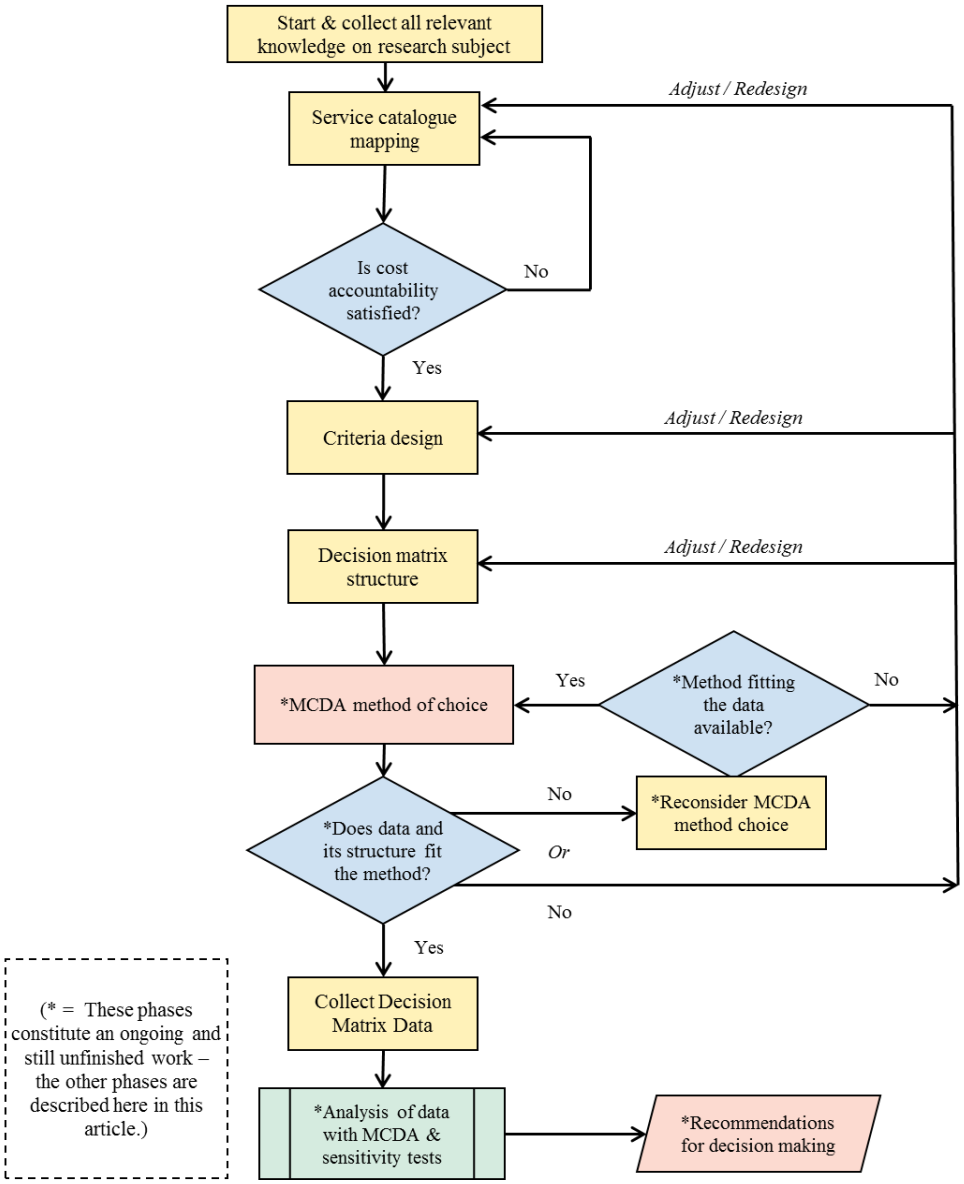
First the research group delved into the general knowledge of the agency as a whole and into its service and function repertoire (see Figure 1, 'service catalogue mapping'). A conclusive list of the agency's services and functions was accumulated in a multi-phase process comprising an interview round in January 2014 (involving experts and decision makers from Trafi and Ministry of Transport), project steering group meetings and e-mail correspondence. Important aspects to be kept in mind throughout this process were

1 to keep the size of the list manageable, and at the same time

2 make sure that in sight of the needs of the later succeeding research phases the cost accountability (or financial allocation) of the committed resources to these services is not lost at any point along the way.

The term 'cost accountability' (see also Figure 1, 'is cost accountability satisfied?') here refers to that later on in the research process there will be a need to pinpoint the exact monetary, human and temporal resources that are committed to and spent in producing the evaluated service, together with data on client payments returning as an income back to the agency upon the delivery of that service. Namely, some services provided by this agency to the general public or other stakeholders are free of charge, but notably others are not – there is a unit price tag on many of the services.

Figure 1 Iterative research design process (see online version for colours)



Note: MCDA = Multiple criteria decision analysis.

Combining points 1 and 2 yields a need to maintain equilibrium between the aggregation level of the list and the cost accountability factor, because otherwise the cost-benefit evaluation later on in the research would be deemed impossible.

In the first try the list was accumulated based on a 'bottom-to-top' approach, by listing the types of services or functions. Each service and function was divided into each transport mode, i.e., aviation, rail, road and maritime. Each specific service was then allocated to these modes. This resulted in a three level categorisation of services, here as examples two isolated lines in the data set:

a Registries -> Road transport -> Road vehicle registry

b Permits and approvals -> Aviation -> Pilot licence.

Later it turned out that the first selected 'bottom-to-top' approach was bound to become challenging in light of the cost accountability requirement described above. Therefore the accumulated list was re-structured, now with a 'top-to-bottom' approach. A classification of targets/objects of services was created (see Figure 2). Within the 'target/function' classification, three of the classes actually clearly refer to a target the services are directed towards, namely 'actor permits and control', 'vehicles' and 'personal permits'. The rest of the classes include the agency's internal functions (or divisions), one top level function per class.

The targets then were classified by transport mode and finally each of the agency's services and functions were allocated to these classifications. Lastly, a service type was attached to each service as meta-data for later use. The service type was divided into six categories of

1 permits and approvals

2 control, monitoring and oversight

3 information service

4 registry service

5 regulation

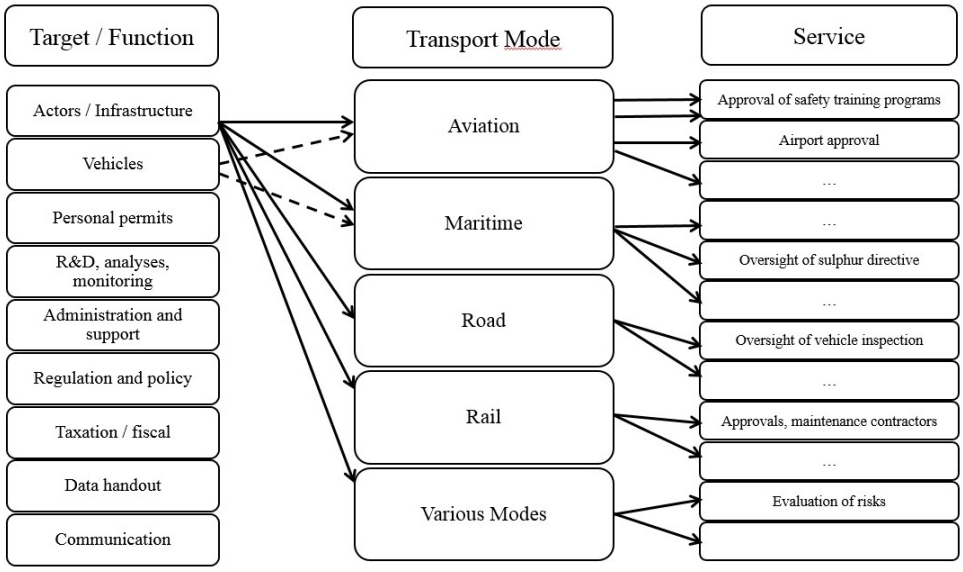
6 other.

The top level classification of the agency's services and functions is illustrated in Figure 2. There the listings for both 'target' and 'mode' are complete but the listing of 'services' is a small sample of the full list.

In addition to the four abovementioned transport modes, a fifth category needed to be added, namely the one called 'various modes'. This category refers to services/functions that affect or deal with more than one transport mode, i.e., deals with two, three or all four modes. An example of such a service is *regulation and policy -> international cooperation* – since the agency's representation in the international bodies for all four transport modes is organised within this one function. Typically also most administrational functions were classified as such, since they do not usually have one single transport mode in focus but instead are more internal in nature and the connection to any specific transport mode is hard or even impossible to trace. This is not to say that these kinds of functions could not have an impact or could not participate in achieving the set goals for the agency - in fact, quite on the contrary. But due to the complex and less than direct connection to any physical activity or induced change on the field (air,

road, rail or sea), these kinds of services cannot be strong candidates to full-scale quantitative economic impact evaluation. Instead, they might well be promising candidates for a more descriptive analysis and a qualitative evaluation.

Figure 2 Classification structure of the service catalogue



The end result of this service cataloguing was a list of 149 cost-accountable services or functions all classified according to the process and structure that is described in the above.

4.2 Selecting the evaluation criteria and scales; pre-processing

The next phase in the research was to select the criteria for the prioritisation exercise (see Figure 1, 'criteria design'). This selection process was based on

1 the general transportation knowledge and impact evaluation research expertise within the research group

2 literature survey.

The agency's main responsibility areas are transportation safety and security, transportation environmental effects, transportation system resiliency, transportation system general status analysis or transportation markets – and furthermore, all of these responsibilities span out across all modes of transport. The research group including the authors of this article together possesses long careers in various fields of transportation research and impact evaluation. This is however not to say that the researchers would be specific experts in all of the fields of the agency's main responsibility areas. Neither is this to say that the research group would have in their possession the best available knowledge on each and every of the agency's services and all aspects of the provision of those services. Here is where the staff of Trafi was needed to step in later: within the

agency namely resides the best available knowledge on all of the abovementioned aspects.

Based on the main responsibility areas, the main criterion sets were formulated. These were divided into three main groups of

1 transportation safety and security

2 transportation environmental effects

3 other criteria.

The selected criterion set was formulated, validated and accepted in close collaboration with the research project steering group and other staff members of the agency, together representing all modes of transport and all main responsibility areas.

For traffic safety (and security) three sub categories were set and under those a total of seven criteria (or impact factors) were selected. Haddon matrix (Peden/World Health Organization, 2004) categorisation was used but in an applied manner. Only the 'x-axis' of the matrix was used, and not the phasing to pre-crash, crash and post-crash stages of a traffic accident. Haddon matrix's x-axis covers well the criteria relating to human factors, vehicle and equipment related factors and environmental factors. The criteria selected for traffic safety (applied from Peden/World Health Organization, 2004) covered the area in the following way: for 'human factors' all passengers, drivers and other personnel for all modes of transport were included; for 'vehicles and equipment' all vehicles and all mobile equipment (such as, e.g., the railway carriages) in all modes were included; and for 'environment' the transport infrastructure in all modes including the associated safety, security and information systems and databases were included.

For transport's environmental impacts, eight criteria were selected. In addition to the accumulated knowledge within the research group, the main basis for the classification of the environmental factors was based into the classification presented in European Environmental Agency (2012). A certain analogue can be detected when comparing the criterion for traffic safety to that of the environment. Namely in that three main criterion sets can be recognised: human factors, vehicle and equipment factors; and transport environment level factors. For other Trafi responsibilities four criteria were defined.

Furthermore, on closer examination and particularly considering the special role and *modus operandi* of the agency and its responsibilities, it was noted that the Haddon matrix factors used in the criteria selection process were lacking one key element, namely the activity in relation to various actors within the transportation field. More precisely in the chains of logics these types of actor related factors are included and mixed in various parts of the framework and consequently in a way remain somewhat hidden to the naked eye – at least hidden more than is appropriate considering the research object. The term 'actors' in regard the agency in question cover a wide variety of establishments such as airport, port and train station owners/operators, logistics operators, public transportation companies, maintenance contractors, vehicle inspection operators, driving schools and other types of education establishments – this list is not exhaustive but is here just to illustrate a selected few types of actors. The agency deals constantly, extensively, directly and on many levels with these actors by, e.g., issuing permits and approvals, auditing their performance, auditing and overseeing their conformity to existing laws and regulations, etc. Because of this key role a supplementary criterion 'actors' was added. In the selection process, this supplementary criterion was added to complement the traffic safety criteria set. The practical reasoning for the need to include the extra criterion

stemmed from an observation that the service catalogue included some services that did not directly tick any of the boxes (or ticked very few of them) in the whole set of criteria especially for traffic safety (to some extent also for environment or other factors) and still the service is seen as a substantial and important part of the agency's activity.

The final set of criteria is presented in Table 1. Note that the term 'individual' includes all relevant groups, i.e., passengers, travellers and other personnel – e.g., in aviation the cabin personnel, the ground personnel at or around airports etc. Similar logic was applied across all transport modes.

Table 1 Evaluation criteria for the agency's mandate areas

Traffic safety and security criteria (CS)		*Environmental criteria (CE)*		*Other criteria (CO)*	
CS1	Individual skills	CE1	Equipment maintenance	CO1	Resiliency, normal circumstances
CS2	Individual condition	CE2	Equipment, propulsion fuel	CO2	Emergency circumstances
CS3	Individual education/attitudes	CE3	Equipment, emissions	CO3	Transportation system status
CS4	Equipment maintenance	CE4	Demand, output	CO4	Transportation markets
CS5	Infrastructure condition	CE5	Infrastructure condition		
CS6	Infrastructure design	CE6	Infrastructure design		
CS7	Infrastructure capacity usage	CE7	Control/overseeing		
CS8	Actor	CE8	Individual mobility choices		

These criteria were then cross-tabulated together with the service catalogue. There each criterion was assigned to each service, each service formed a row and each criterion formed a column in the table. This resulted into a 149 times 20 tables (149 services, 20 criteria, a total of 2,980 cells), which later will be referred to as the 'decision matrix' (Belton and Stewart, 2002) – albeit at this point an empty draft shell in regard the cell values. In some literature the decision matrix is also called the 'performance matrix' (Figueira et al., 2013) (see also Figure 1, 'decision matrix structure').

Next the scale for the evaluation was designed. A five point scale was selected. The purpose of the prioritisation exercise was not at this point to evaluate or form opinions about the actual impact volumes of any services, but instead to narrow down the list of services to those that the actual impact evaluation later would be performed to. Therefore upon completing the cell values within the decision matrix, the question posed to the decision maker was "does the service under consideration have an imaginable impact in regard the criterion under consideration". If the answer to the question was 'no', the cell was to receive a null value. If the answer was 'yes', a further complementing question was posed, namely "how obvious and/or direct the consequence mechanism is?" Based on the answer to that question, the cell was allocated a value from 1 to 4 using the following scale:

- 4: extremely obvious and/or direct consequence relation exists
- 3: obvious and/or direct consequence relation exists

- 2: an imaginable consequence relation exists
- 1: there is a consequence relation, but it is very indirect, complex or obscure
- (0: no imaginable consequence relation).

The selected scale is discrete and not true or absolute. It is a subjectively defined five point ordinal scale designed to leave room for the decision makers' specific knowledge, preferences and opinions to come through.

Subsequently, the research group conducted a pre-evaluation of the decision matrix, using the abovementioned scale. This in practice meant pre-filling most of the cell values based on all of the available information and the research group's expertise. The main reasoning behind having the matrix pre-filled was to make the data collection workshop much more manageable both in size and in duration. In cases where the research group was even slightly insecure or unsure either of what the service content actually is or what (if any) cause-consequence mechanisms could exist in relation to any particular criteria, the cell was automatically and strictly left blank.

Large proportion of the cells was given a definite null value. Namely, in most cases, a service has been designed and directed to tackle a certain aspect of the agency's main responsibilities (or policy targets) – and not all of them. An example to illustrate this: a service that is designed to and aimed at controlling the validity or quality of driver education very seldom has imaginable impacts or implications in regard infrastructure design aspects (note: word 'seldom' instead of 'never' used since although the presented example is straightforward, all cases were not). Therefore to avoid oversight bias the researchers needed by default to adopt openness to consideration of 'not necessarily never'. Again, whenever in any doubt, cell was to be left blank). Two examples to illustrate the valuation logic are presented below. Both examples present one criteria and then two Trafi services valuated in regard that criteria.

> Example 1: Traffic safety criteria 'Driver education and attitudes', the question to be considered is "Does the service under consideration have an imaginable impact into the quality of drivers' education or knowledge levels?" Trafi service 'Railway education organisation permits' most probably has an impact, so the cell should receive a value between 1 and 4. Trafi service 'Aircraft registries' does not have an impact, so the cell receives a null value.
>
> Example 2: Environmental criteria 'propulsion fuel', the question to be considered: "Does the service under consideration have an imaginable impact into the vehicle buyers' choices in regard what propulsion fuel the new vehicle will use (i.e., gasoline, diesel, electric, fuel cell, hybrid, etc.)?" Trafi service 'Vehicle taxation' most probably has an impact on the choices, so the cell value should be 1–4. Trafi service 'Aviation education permits' does not have an impact, so the cell receives a null value.

4.3 Compiling and completing the decision matrix

Next all pre-valuated data (pre-filled decision matrix) was completed (empty cells), corrected (pre-filled cells) and validated (all cells) in two phases. Firstly a one day workshop was arranged to tackle the main bulk of this workload. Secondly, some elements of the matrix went through another checking, evaluation and scrutiny round

within the agency. The 14 Trafi employees participated in the workshop, and additionally two research group members were present in supporting roles. Supporting role means that the researchers provided the guidance for the working methods, distributed the instructions and material, and operated as moderators/facilitators of the workshop – but did not take part in any way into actually valuating the services. The background of the participating Trafi personnel was designed and invited so that all relevant aspects were covered: all transport modes were represented; the group included experts both from service provision layer and company top level management; and experts for all of the agency's responsibilities were there (e.g., transport safety, environment, etc.). During the workshop it turned out that the representation for both aviation service expertise and expertise on services covering more than one transport mode were slightly inadequate – thus a secondary off-line scrutiny round was introduced. The service group 'various modes' was observed to include several strategically important functions (such as the international cooperation activities, functions dealing with issuing norms and regulations or supporting national legislation development) that deserved to be separately evaluated by representation from each transport mode.

4.4 Running the workshop and the actions thereof

Four groups were formed so that aviation, maritime and road modes had one group each and additionally the fourth group included rail transport services and additionally the services concerning more than one transport modes (see centre column of Figure 2 for reference). The logic and reasoning behind deciding to combine 'rail related services' with 'various modes' was based simply on the considerably smaller number of rail related services in comparison to other transport modes – thus the aim was to even out the work load between the groups. Later on in the research however, both 'rail' and 'various modes' will be addressed as separate 'modes'.

Each group was allocated at least one representative from the project steering group would be present to provide, if needed, further in-depth insight for the rest of the group about the project and other background information. One or more representatives from the associated transport mode were allocated to each group, and a representative from the more practically oriented service provision layer was allocated correspondingly.

Workshop material included A0 size printouts of the service catalogues relevant for each group, post-it notes, notepads, writing equipment and a sheet highlighting the important things to be kept in mind during the session. This sheet included a reminder of the required mind-set and viewpoint for the group's task as described under 'the evaluation scale' above, and an additional important reminder to constantly refrain from trying to answer questions like "how much does this service impact this or that criterion/factor?" or "how substantial this service is in terms of its production volume or allocated resources?".

During the group work phase, the facilitators constantly circled around and physically visited each group several times mainly to monitor that all the agreed instructions were followed to the letter. They only intervened where direct questions were posed and in cases where they detected that the discussion within the group was at risk of derailing from the instructed objective and path, or from the limited schedule.

4.5 Discussion and conclusion session

The group work results were collated into one document and a first look was taken into what was achieved. Groups 'aviation services' and the latter part of 'rail services and services concerning various modes' were re-checked following the same procedure and mind-set as in the group work phase. After this the complete data sets were provided to the research group.

5 Decision matrix to facilitate evaluation

The above described iterative phases of service catalogue mapping, criteria design, pre-filling and evaluation within the workshop, the end result was yielding the completed decision matrix. This is the data for the next research phase of decision analysis to prioritise services for further analysis. After the process, the decision matrix at this point included 149 services valuated on a scale from 0 to 4, each in regard 20 criteria. Inherently, all transport modes reside in the matrix with indifferent weighting in comparison to any other transport mode group. This means that no mode is deemed more important than the other, regardless of the manner of how generously or sparingly the group ended up giving their valuation ratings on the 0–4 scale. The strength of this approach is that the most promising and potential services to be evaluated in full-scale in the further research will be highlighted separately for each mode category. Failing to do so would risk ending up in a lopsided and hence unsatisfactory sample of the agency's service and responsibility portfolio. Table 1 illustrates the decision matrix format and contents, where a small sample of seven aviation related services and their decision matrix valuations, only in regard traffic safety and security, are presented (in other words this is only a small 7 × 8 excerpt from the full 149 × 20 DM, for the purpose of showing an example). Just looking at one of the services, one can observe that 'organisation permit, [aviation] personnel education' has been evaluated with maximum score for criteria CS1 'driver (or passenger or traveller or other personnel) skills', and correspondingly, with a minimum score for criteria CS6 'infrastructure design'.

Table 2 Decision matrix (excerpt) for aviation related services in regard traffic safety and security criteria

Aviation related services	*Traffic safety and security criteria*							
	*CS1**	*CS2*	*CS3*	*CS4*	*CS5*	*CS6*	*CS7*	*CS8*
Control of personnel permits	4	2	3	3	1	0	0	2
Aviation operator permit	3	1	2	3	1	1	0	4
Terminal/airport level inspections	1	3	2	3	2	0	1	3
Organisation permit, personnel education	4	1	3	2	2	0	0	2
Overseeing permits for organisations	2	2	3	2	1	0	0	4
Maintenance personnel certificates	4	1	3	3	0	0	0	2

Note: *See Table 1 for the abbreviations.

The data accumulated with the procedure will be used for multi criteria decision analysis (see Figure 1, 'analysis of data with MCDA and sensitivity tests'). This analysis will yield the recommendations for the research sample selection. Based on very initial MCDA analysis, the services that proactively affect traffic safety and environment seem to gain quite a lot attention. This could implicate that Trafi's risk-reducing type of services are those that will be emphasised even more in the future. However, this finding is still preliminary.

6 Discussion and sum-up

The need to evaluate benefit-cost ratios or cost-effectiveness of Trafi's services stem from wider evolution and change processes of administration. The change processes were started already in the 1980's as part of global transition towards new public governance and administration models that adopt managerial tools and practices merging from the business sector. The evolution history of Trafi shows that these processes of institutional redesigns are ongoing.

From Trafi's management point of view, the tools that demonstrate the benefits and effectiveness of Trafi's services are of crucial importance. On one side, demonstrating efficiency and effectiveness as well as wider socio-economic impacts justifies the whole existence of the administration's current structure and service architecture. On the other side, the internal budget competition between public administration sectors leads to increasing need of proving positive impacts. Hence both the upper level administration – i.e., the Ministry – and Trafi's own management cannot afford not to show the case for Trafi and its service repertoire.

This paper showed the first steps in a research process for prioritising the services that should be evaluated in more detail. It goes without saying that not all services are even possible to be evaluated and some other services, mostly based on existing legislation, are a must in an administration's service architecture. However, one should also emphasise that excessive metrics in management control could lead to controversial results (see e.g., Vakkuri, 2013; Martin and Mikovsky, 2010) and reduce agility. Furthermore, it could lead to reduced accountability, meaning that accountability may be reduced to those issues covered by performance metrics only.

The results achieved by this research indicate that collection of data for prioritisation of services is possible in a structured manner. The selected approach suited well for the data collection, and in later stages it was observed that the data was well suited for use within an MCDA method. Furthermore it was observed that the process presented is highly normative, so similar approaches can be implemented elsewhere with ease. It should be noted, though, that any prioritisation is dependent on reference groups' preferences and perceptions and by definition are coloured according to the interests of the group. Therefore, prioritisation could also be considered to involve several interest groups in order to gain multiple angles and engage a wider community to assess which services are essential, relevant and meaningful to be included in the impact evaluation. For this particular study (and its perspective and needs) though, both the selected interest group and the approach were appropriate.

Another challenge to be dealt with in future research is the chronological preference. In other words, some services are bound to be more relevant in the long run and hence

carry a strategic loading. For example, the services or tasks related to international cooperation with European Commission and EU central agencies are time consuming but the prospective impacts of lobbying and influence work could be substantial in the long run. Examples of directives, for example, that have radical impact on transport sector's performance are known. For instance, the new regulation on sulphur content in marine fuels (European Union, 2012) seem to result in unforeseen effects on Finland's commercial marine operators and shipping companies.

The next step (ongoing at the time of writing this) in the research is the application of the acquired decision matrix within applicable MCDA methods. Some limitations apply in further use of the approach as presented in this article. The findings presented here are discoveries from the material that was used and therefore they cannot be directly generalised. The analysis describes the setup as it was embedded in the local context. However, repeating the approach in other settings, such as for corresponding agencies in other countries, would be highly interesting. In that way further information on the versatility and generalisability of the presented approach could be acquired. The developed process proved successful, hence clearing path for similar analyses on other agencies, administrations and services elsewhere. Even more sophisticated methods are hopefully being considered, but the approach described here is one potential foundation for future development.

Acknowledgements

The authors would like to acknowledge the input of Professor Josef Jablonský, University of Economics Prague, during the early stages of this research.

References

Belton, V. and Stewart, T.J. (2002) *Multiple Criteria Decision Analysis: An Integrated Approach*, Kluwer Academic Publishers, London.

European Commission (2010) 'Towards a European road safety area: policy orientations on road safety 2011–2020', *COM (2010)*, 389 final, Brussels.

European Commission (2011) 'Impact assessment', *SEC (2011)*, 358 final, Commission Staff Working paper, Brussels.

European Environment Agency/COWI (2013) *Evaluation of the European Environment Agency*, final report, COWI A/S, Denmark.

European Environmental Agency (2012) *TERM 2012: Transport Indicators Tracking Progress Towards Environmental Targets in Europe*, European Environmental Agency Report 10/2012, Brussels.

European Union (2012) 'Directive 2012/33/EU of the European Parliament and the Council of 21 November 2012 amending Council Directive 1999/32/EC as regards the sulphur content of marine fuels', *Official Journal of the European Union Brussels*, Vol. L 327, pp.1–13.

Figueira, J.R., Greco, S.B., Roy, B.C., Słowiński, R. (2013) 'An overview of ELECTRE methods and their recent extensions', *Journal of Multi-Criteria Decision Analysis*, January, Vol. 20, Nos. 1–2, pp.61–85.

Finnish Ministry of Finance (2012) *Vaikuttavuus ja tuloksellisuusohjelma 2012 [National Program for Impacts and Efficiency Guideline]*, Finnish Ministry of Finance, Helsinki.

Finnish Ministry of Justice (2009) *Laki ja asetus liikenteen turvallisuusvirastosta (863/13.11.2009 ja/ and 865/13.11.2009) [Finnish Law and Statute on Finnish Traffic Safety Agency]*, Finnish Ministry of Justice, Helsinki.

Finnish National Audit Office VTV (2013) *Tuloksellisuustarkastuskertomus.Liikenneturvallisuus [Audit of national traffic safety administration performance]*, Audit report 6/2013, Finnish National Audit Office, Helsinki.

Fuller, D., Gauvin, L., Kestens, Y., Daniel, M., Fournier, M., Morency, P., Drouin, L. (2013) 'Impact evaluation of a public bicycle share program on cycling: a case example of BIXI in Montreal', *American Journal of Public Health*, Vol. 103, No. 3, pp.e85–e92, Quebec.

Gruening, G. (2001) 'Origin and theoretical basis of new public management', *International Public Management Journal*, Vol. 4, No. 1, pp.1–25.

Haque, M.S. (2004) 'New public management: origins, dimensions, and critical implications', in Tummala, K.K. (Ed.): *Public Administration and Public Policy*, Eolss Publishers Ltd., Oxford UK.

Martin, L.L. and Mikovsky, L.P. (2010) 'Comparative performance measurement in decentralised systems: the case of the Florida Benchmarking Consortium', *International Journal of Public Sector Performance Management*, Vol. 1, No. 4, pp.376–389.

Maugeri, S. and Metzger, J-L. (2013) 'Public action: a question of performance?', *International Journal of Public Sector Performance Management*, Vol. 2, No. 2, pp.105–122.

Ministry of Transport and Communications Finland (MTCF) (2012) *Liikenne – ja viestintäministeriön ja Liikenteen turvallisuusviraston välinen tulossopimus vuosille 2013–2015. [Wage settlement between MTCF and Trafi for years 2013–2015]*, Ministry of Transport and Communications Finland, Helsinki.

Ministry of Transport and Communications Finland (MTCF) (2013) *Tieliikenteen turvallisuustoiminnan työnjako, selvitysmiehen raportti [The division of traffic safety work, rapporteur Dr.Matti Roine]*, Publications 29/2013, MTCF, Helsinki.

Mononen, P., Leviäkangas, P., Roine, M. and Haapasalo, H. (2014) *Trafi and the Surrounding Transport Policy Framework*, University of Oulu Working papers in IEM 2/2014.

OECD (2010) *OECD Public Governance Reviews, Finland, Working Together to Sustain Success, Assessment and Recommendations*, OECD, Paris, France.

Peden/World Health Organization (2004) in Margie (Ed.): *World Report on Road Traffic Injury Prevention*, World Health Organization, Geneva.

Proost, S., Dunkerley, F., Van der Loo, S., Adler, N., Bröcker, J. and Korzhenevych, A. (2014) 'Do the selected Trans European transport investments pass the cost benefit test?', *Transportation*, Vol. 41, No. 1, pp.107–132, Springer, Netherlands.

Trafi (2012) *Annual Report 2011*; TRAFI/1811/02.00.00/2012, Trafi, Helsinki.

Vakkuri, J. (2013) 'Interpretive schemes in performance management – measurement problems generating managerial action in Finnish local government', *International Journal of Public Sector Performance Management*, Vol. 2, No. 2, pp.156–174.